Polymorphisms at the innate immune receptor TLR2 are associated with borrelia infection in a wild rodent population

Tschirren, Barbara ; Andersson, Martin ; Scherman, Kristin ; Westerdahl, Helena ; Mittl, Peer R E ; Raberg, Lars

Abstract: The discovery of the key role of Toll-like receptors (TLRs) in initiating innate immune responses and modulating adaptive immunity has revolutionised our understanding of vertebrate defence against pathogens. Yet, despite their central role in pathogen recognition and defence initiation, there is little information on how variation in TLRs influences disease susceptibility in natural populations. Here we assessed the extent of naturally occurring polymorphisms at TLR2 in wild bank voles (Myodes glareolus) and tested for associations between TLR2 variants and infection with Borrelia afzelii, a common tick-transmitted pathogen in rodents and one of the causative agents of human Lyme disease. Bank voles in our population had 15 different TLR2 haplotypes (ten different haplotypes at the amino acid level), which grouped in three well-separated clusters. In a large-scale capture-mark-recapture study we show that voles carrying TLR2 haplotypes of one particular cluster (TLR2c2) were almost three times less likely to be Borrelia-infected than animals carrying other haplotypes. Moreover, neutrality tests suggested that TLR2 has been under positive selection. This is the first demonstration of an association between TLR polymorphism and parasitism in wildlife, and a striking example that genetic variation at innate immune receptors can have a large impact on host resistance.

DOI: https://doi.org/10.1098/rspb.2013.0364

Polymorphisms at the Innate Immune Receptor *TLR2* are Associated with *Borrelia* Infection in a Wild Rodent Population

Barbara Tschirren[1*], Martin Andersson[2], Kristin Scherman[2], Helena Westerdahl[2], Peer R. E. Mittl[3] and Lars Råberg[2]

[1]Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

[2]Molecular Ecology and Evolution Lab, Department of Biology, Lund University, Sölvegatan 37, SE-223 62 Lund, Sweden

[3]Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Short title: *TLR2* Mediates *Borrelia* Resistance

Keywords: wildlife disease, host-parasite interactions, Borrelia, innate immune defence, Toll-like receptors, disease resistance

*Correspondence: Barbara Tschirren, Institute of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland; Tel: +41 44 635 47 77, Fax: +41 44 635 68 18; Email: barbara.tschirren@ieu.uzh.ch

25 Abstract

26 The discovery of the key role of Toll-like receptors (TLRs) in initiating innate
27 immune responses and modulating adaptive immunity has revolutionised our
28 understanding of vertebrate defence against pathogens. Yet, despite their
29 central role in pathogen recognition and defence initiation, there is little
30 information on how variation in *TLRs* influences disease susceptibility in
31 natural populations. Here we assessed the extent of naturally occurring
32 polymorphisms at *TLR2* in wild bank voles (*Myodes glareolus*) and tested for
33 associations between *TLR2* variants and infection with *Borrelia afzelii,* a
34 common tick-transmitted pathogen in rodents and one of the causative agents
35 of human Lyme disease. Bank voles in our population had 15 different *TLR2*
36 haplotypes (ten different haplotypes at the amino acid level), which grouped in
37 three well-separated clusters. In a large-scale capture-mark-recapture study
38 we show that voles carrying *TLR2* haplotypes of one particular cluster
39 ($TLR2_{c2}$) were almost three times less likely to be *Borrelia*-infected than
40 animals carrying other haplotypes. Moreover, neutrality tests suggested that
41 *TLR2* has been under positive selection. This is the first demonstration of an
42 association between *TLR* polymorphism and parasitism in wildlife, and a
43 striking example that genetic variation at innate immune receptors can have a
44 large impact on host resistance.

Introduction

Parasites, by definition, are harmful to their hosts and should therefore impose selection for enhanced resistance. Despite this, there is typically significant genetic variation for resistance to parasites and pathogens in natural populations [1]. To elucidate the evolutionary causes and consequences of such variation, a better understanding of which genes actually contribute to variation in resistance is desirable [2-4]. However, whereas there is a considerable body of literature on the genetic basis of resistance in humans [5], knowledge from other animals, and in particular from natural vertebrate populations, is as yet very limited (but see [6]).

The principal weapon that hosts have evolved to fight off pathogens is the immune system, which in vertebrates consists of two main parts, innate and acquired immunity [7]. Invading infectious agents are initially recognized by the innate branch of the immune system through pattern-recognition receptors (PRRs). Pattern-recognition receptors recognize structures that are specific to microbes (MAMPs or PAMPs; microbe- / pathogen-associated molecular patterns). After stimulation by the ligand, PRRs activate an intracellular signaling cascade, which initiates innate and acquired immune responses [8-10]. An important class of PRRs are Toll-like receptors (TLRs), which were discovered in vertebrates as late as 1997 [11]. Most mammals have 10–13 different TLRs, each recognizing different ligands [12]. TLRs have been found to play a key role in pathogen recognition and initiation of immune responses in humans and laboratory animals [13, 14], and there is an increasing number of studies in humans showing associations between TLR polymorphisms and infectious diseases [14-16]. Yet,TLR polymorphisms have, unlike

70 polymorphisms at the Major Histocompatibility Complex (MHC) [17], thus far
71 received little attention from ecologists investigating host-parasite interactions
72 and wildlife disease (but see e.g. [6, 18, 19]).
73 Here we investigated the role of naturally occurring *TLR2*
74 polymorphisms in mediating parasite resistance (here defined as the ability to
75 prevent and / or clear infection, and measured as the presence / absence of
76 infection) in a population of wild-living bank voles (*Myodes glareolus*) by
77 testing for associations between *TLR2* genotype and *B*orrelia *afzelii* infection
78 status. *Borrelia afzelii* is a common tick-transmitted pathogen in rodents [20],
79 and one of the causative agents of human Lyme borreliosis, the most
80 common zoonotic disease in Europe and North America [21, 22].
81 Lipopeptides, which are central components of the cell walls of *Borrelia*, are
82 ligands for TLR2, and knock-out studies with laboratory mice have shown that
83 TLR2 plays an important role in the recognition and initiation of immune
84 responses against *Borrelia* [23-27]. Moreover, there is evidence that a
85 common single nucleotide polymorphism (SNP) in the human *TLR2* affects
86 susceptibility to Lyme disease [28]. By transferring this immunological
87 background knowledge into an ecological context, we here show that
88 polymorphisms at *TLR2* are associated with *B. afzelii* infection in a natural
89 rodent population, highlighting the important role of Toll-like receptors in
90 mediating disease susceptibility in wildlife.
91
92
93
94

95 Material and Methods

96

97 *Study species*

98 The bank vole (*Myodes glareolus,* Rodentia*)* is one of the main hosts of
99 *Borrelia afzelii* in Europe. *Borrelia afzelii* is transmitted by the sheep tick
100 (*Ixodes ricinus*) between hosts [29]. We captured bank voles in 2008 in Kalvs
101 mosse (N 55° 42.470', E 13° 29.216'), a homogenous, deciduous woodland of
102 about 0.25 km^2 south-east of Revingeby, Skåne, Southern Sweden using live-
103 traps (Ugglan Special No1, Grahnab, Gnosjö, Sweden). Animals were caught
104 during trapping sessions in May (N=31; 100% adults), June (N=171; 45%
105 adults), August (N=252; 43% adults), September (N=320; 29% adults), and
106 October (N=350; 29% adults). They were weighed (± 0.1g), and the number of
107 tick larvae on the ears was counted as a proxy for infestation with nymphs,
108 the main infective stage [30]. Molecular sexing was performed by amplifying a
109 fragment of the male-specific sex-determining region Y (SRY) as described in
110 Wandeler et al (2008) [31]. At first capture, animals were individually marked
111 with subcutaneous transponder tags (Trovan ID-100B, AEG ID, Ulm,
112 Germany) to allow for the identification of individuals upon recapture. We
113 obtained ear biopsies from 726 individuals during the trapping sessions.
114 Samples were stored in 70% ethanol for later DNA extraction (as described in
115 [19, 32]), *TLR2* genotyping and determination of *Borrelia* infection status. All
116 animal procedures were performed under licences M101-06 and M141-10
117 issued by the Malmö / Lund, Sweden ethical board for animal experiments.

118

119 *Toll-like receptor 2 (*TLR2*) genotyping*

120 Mammals have a single functional copy of *TLR2* [12]. There is a *TLR2*
121 pseudogene in humans and dogs, but not in mice [12]. In bank voles, the
122 entire *TLR2* coding region is 2352 bp long [19]. For this study, we sequenced
123 a 1173 bp long fragment of *TLR2* from bp 691 to 1863 as described in
124 Tschirren et al (2012) [33]. There was no indication that we amplified more
125 than one locus (at most two peaks per site in the chromatograms), and we
126 found no sign of pseudogenes (no stop codons or frame shift mutations). The
127 sequenced part of *TLR2* contains most of the functionally relevant sites
128 involved in pathogen-recognition and *TLR2-TLR1* heterodimerisation [34, 35],
129 and we previously demonstrated molecular signatures of positive selection
130 during the evolutionary history of rodents [19], as well as strong population
131 differentiation and isolation by distance across bank vole populations within
132 this gene region [33]. The amplicon consisted of coding sequence only.
133 Sequences were processed, assembled and aligned using Geneious 5.0.4.
134 [36], and all polymorphisms were examined by eye. *TLR2* haplotypes were
135 reconstructed with PHASE 2.1 [37, 38] using the default settings of a thinning
136 interval of one, 100 burn-in iterations and 100 main iterations. Haplotypes
137 were submitted to NCBI GenBank (see Table 1 for Accession numbers).
138 To test if patterns of haplotype frequencies and tree topology are consistent
139 with neutral expectations, we performed two neutrality tests: *Fay and Wu's H*
140 [39] and *Li's MFDM* [40]. For the former, the empirical distribution of the test
141 statistics was generated using neutral coalescent simulations in DNAsp [41],
142 based on the observed number of segregating sites, 20'000 replicates and the
143 assumption of no recombination (no recombination was detected with the
144 *MFDM* test). The results of this analysis did not change qualitatively when

allowing for moderate levels of recombination (data not shown). For both tests, we used the *Mus musculus* (NM_011905.3), *Apodemus flavicollis* (JN674549.1) and *Myodes rufocanus* (HM215593.1) *TLR2* sequences as outgroups. Deviation from neutrality detected by *Fay and Wu's H* can be caused by selection and / or demographic processes such as population expansion or bottlenecks [40, 42]. Li's *MFDM* is robust against population size changes, but can be sensitive to admixture events [40].

To identify clusters of host haplotypes we constructed a *TLR2* haplotype network in TCS 1.21 [43]. Homology models for the most common haplotype of each cluster (haplotype 1 and 6, see Results) were generated based on the human TLR1-TLR2 lipopetide crystal structure [35]. Alignments were generated with ClustalW [44], manually curated where necessary, and served as the input for program Modeler 9v4 [45]. Figures were prepared using PyMol (http://www.pymol.org/).

Borrelia *infection status*

To determine if voles were infected with *B. afzelii*, we performed *flaB* real-time PCR assays as described in Råberg (2012) [46] (*N* = 1124 samples). Samples with a melting temperature between 78.15 and 78.75 °C and a Ct value corresponding to ≥1 *B. afzelii* spirochaete were considered positive. The PCR assay is specific for *B. afzelii* [46].

Borrelia afzelii is the only Lyme borreliosis-causing *Borrelia species* observed in our bank vole population. The relapsing fever-causing *Borrelia species B. miyamotoi* is also found, although at very low prevalence [46].

169 Whereas *B. afzelii* infection status is highly repeatable (once infected, an
170 individual stays infected for life [47]), we found that an individual's infection
171 intensity (number of *B. afzelii* spirochaetes per unit host tissue) varied
172 considerably over time. Therefore, we focused on infection status only.
173
174 *Statistical analyses*
175 The reconstructed *TLR2* haplotype network revealed that *TLR2* haplotypes
176 grouped in three clearly separated clusters, one of which was very rare (see
177 Results). To increase the statistical power to detect differences in *B. afzelii*
178 infection among individuals with functionally different *TLR2*, we focused on
179 the two common clusters and determined for each individual bank vole if it
180 carried two *TLR2* haplotypes belonging to the first cluster ($TLR2_{c1}$), two
181 haplotypes belonging to the second cluster ($TLR2_{c2}$), or one haplotype of
182 each cluster (see Fig. 1, ESM 1). This approach is very powerful, yet
183 conservative because it assumes that haplotypes within clusters are
184 functionally identical. We used a generalised linear mixed model with a
185 binomial error structure to test for differences in *B. afzelii* infection between
186 host *TLR2* clusters. We included number of haplotypes belonging to cluster
187 $TLR2_{c2}$ (zero, one, or two), host age class (adult: >20g, subadult: 15-20g,
188 juvenile: <15g, [48]), sex and their two-way interactions as fixed effects in the
189 statistical model. Individual ID and trapping session (month) were included as
190 random effects to account for the non-independence of measures from an
191 individual captured during more than one trapping session and for seasonal
192 variation in *Borrelia* prevalence.

193 Because differences in *B. afzelii* infection among individuals could be due to
194 differences in resistance or exposure, we ran the same model, but with a
195 poisson error structure, to test for differences in tick load (i.e. the *Borrelia*
196 vector) between *TLR2* clusters. We used number of larvae as a proxy for
197 exposure to nymphs (the main infective stage), because nymphs are
198 comparably rare and therefore difficult to quantify accurately [30].
199 Analyses were run in R 2.14.1 [49] using the glmer function, part of the lme4
200 package [50]. For all analyses, the significance of the fixed effects was
201 determined by comparing two nested models, with and without the factor of
202 interest, using likelihood-ratio tests.
203 In addition, we also used a model selection procedure using AIC_c criteria to
204 determine which model best explains variation in *Borrelia* infection status.
205 Candidate models contained combinations of *TLR2* genotype, age class and
206 sex. All candidate models contained individual ID and trapping session as
207 random effects. Model selection was performed using the MuMIn package in
208 R 2.14.1 [49].
209
210
211 Results
212 TLR2 *diversity in wild-living bank voles*
213 *TLR2* diversity was high in the surveyed bank vole population with 15 unique
214 DNA haplotypes, of which ten differed at the amino acid sequence. The most
215 common amino acid haplotype (haplotype 1) occurred in six variants ($1_{a\text{-}f}$),
216 which differed at the nucleotide, but not the amino acid level (i.e. synonymous
217 substitutions only). No synonymous variants were observed in the other nine

haplotypes. The frequencies of the different *TLR2* haplotypes are shown in Table 1 (N= 726 individuals).

Both neutrality tests indicated that positive selection has shaped *TLR2*. Fay and Wu's test detected an excess of high-frequency derived haplotypes in the population (*Mus musculus* as outgroup: H = -14.95, $P < 0.001$; *Apodemus falvicollis* as an outgroup: H = -15.21, $P = 0.012$); *Myodes rufocanus* as outgroup: H = -13.27, $P = 0.023$). Similarly, *Li's MFDM* test, which uses tree topology to infer selection, was significant ($P < 0.009$).

A reconstructed haplotype network revealed two major *TLR2* clusters ($TLR2_{c1}$ *and* $TLR2_{c2}$; Fig. 1). The main difference between the two clusters were six linked, nonsynonymous SNPs spread out over a region of 290 amino acids (in leucine-rich repeat (LRR) 10 – C-terminal LRR domain; [35], ESM 1). Cluster $TLR2_{c1}$ consisted of haplotypes 1, 2, 3, 4, 14 (63.3% overall frequency) and was defined by the six-site amino acid combination '276Thr/417Asp/453Met/484Ile/536Val/565Asn', whereas cluster $TLR2_{c2}$ consisted of haplotypes 6, 7, 8, 9 (33.4% overall frequency, Fig. 1) and was defined by the alternative amino acid combination '276Ala/417Gly/453Thr/484Leu/536Ile/565Asp'. One rare haplotype (10, frequency 3.3%; Fig. 1) did not group with either of the two major clusters, but formed a third, well-separated group (Fig. 1). We did not consider this third cluster in the analyses because of its low frequency. Because the six high frequency SNPs that separated the two major haplotype clusters always co-occurred (i.e. were perfectly linked), we tested for associations between *Borrelia* infection *and TLR2* clusters rather than individual SNPs in the subsequent analyses. *TLR2* SNPs that separated haplotypes *within* clusters

were not considered in the analysis because they were relatively rare (Table 1; frequency of homozygotes < 3%) and statistical power to detect associations between *Borrelia* infection and these SNPs was therefore low.

Structural differences between TLR2 *haplotypes*

To investigate if any of the six linked SNPs that separated the two major haplotype clusters could potentially affect ligand binding, we modelled the structure of the bank vole TLR2. Based on the human TLR2-TLR1 lipopeptide crystal structure [35] we generated homology models for the most common *TLR2* haplotype of each cluster, haplotype 1 and 6. These models confirmed a high degree of structural conservation between the human and bank vole TLR2-TLR1 complex (55 % sequence identity for *TLR2* residues 206-549). Furthermore, all residues that are involved in the TLR1-TLR2 interface are conserved between humans and bank voles, suggesting that human and bank vole TLR2 possess similar modes of action. None of the polymorphic sites are located directly in the TLR1-TLR2 interface, but the polymorphic site at amino acid position 276 in LRR 10 (ESM 1) is likely to have a significant impact on ligand binding. The side chain of the amino acid at position 276 is pointing into the hydrophobic core of the TLR2-TLR1 complex, and is located in close proximity (10.3 Å) to the putative binding site for the *Borrelia* lipopeptide (Fig. 2). The amino acids found at this position (276Ile in humans) in the two haplotype clusters (276Thr vs. 276Ala) differ markedly in size and polarity, which, given their location, is likely to affect the size of the hydrophobic pocket, and thereby ligand binding. The polymorphic sites 417, 453 and 484 in the TLR2-TLR1 complex are located more than 18 Å away from the ligand-

268 binding site, and the side chains of residues at positions 417 and 453 are
269 pointing towards the solvent, which makes it unlikely that they influence
270 ligand-binding directly. However, they could still affect the thermodynamic
271 stability of TLR2. Polymorphic sites 536 and 565 were outside the template
272 structure (human TLR2-TLR1 complex; [35]) and could not be modelled.
273 However, because of their location in LRR20 and the C-terminal domain we
274 would not expect a prominent effect of these amino acid mutations on ligand
275 affinity [35].
276
277 TLR2 *polymorphisms are associated with* Borrelia *infection status*
278 Overall prevalence of *B. afzelii* infection reached 34.1% in adult bank voles
279 (>20g), but was markedly lower in subadults (15-20g; 10% infected) and
280 juveniles (<15g; 5.6% infected; $\chi^2 = 99.38$, $DF = 2$, $P < 0.001$). Similar results
281 were obtained when analysing differences in *Borrelia* prevalence across age
282 classes for each trapping session separately (ESM 2). The low *Borrelia*
283 prevalence in juveniles and subadults is likely due to limited *Borrelia* exposure
284 rather than higher resistance. Indeed, subadults and juveniles were infested
285 with > four-times fewer ticks than adult voles ($\chi^2 = 92.67$, $DF = 2$, $P < 0.001$).
286 Again, similar results were obtained when analysing differences in tick load
287 across age classes for each trapping session separately (ESM 3). Because of
288 differential *Borrelia* exposure, associations between genetic determinants of
289 resistance and *Borrelia* prevalence are predicted to be pronounced in adults,
290 but much weaker, or absent, in juveniles and subadults. In line with this
291 prediction, there was a significant interaction effect between age class and
292 *TLR2* genotype on *Borrelia* infection status ($\chi^2 = 10.26$, $DF = 4$, $P = 0.036$). No

293 significant association between *TLR2* clusters and *Borrelia* infection was
294 observed in juveniles (χ^2 = 2.55, *DF* = 2, *P* = 0.279) or subadults (χ^2 = 0.24,
295 *DF* = 2, *P* = 0.886). Adult bank voles, however, differed significantly in *Borrelia*
296 infection status depending on their *TLR2* genotype (χ^2 = 12.27, *DF* = 2, *P* =
297 0.002, Fig. 3). *Borrelia* prevalence was lowest in adult voles carrying two
298 haplotypes belonging to cluster $TLR2_{c2}$ and highest in individuals carrying two
299 haplotypes belonging to cluster $TLR2_{c1}$. Voles with one haplotype of each
300 cluster had intermediate infection prevalence (Fig. 3). A posteriori tests
301 revealed that individuals carrying one haplotype of each cluster were
302 significantly less likely to be *Borrelia* infected than individuals with two
303 haplotypes of cluster $TLR2_{c1}$ (χ^2= 5.99, *DF*=2, *P* = 0.014). However, they were
304 not significantly different from individuals with two haplotypes of cluster
305 $TLR2_{c2}$ (χ^2= 2.39, *DF*=2, *P* = 0.122), likely due to a large confidence interval in
306 this latter group (Fig. 3).
307 Males were more likely to be *Borrelia* infected than females (χ^2= 4.51, *DF*=2,
308 *P* = 0.034), but the association between *TLR2* clusters and *Borrelia* infection
309 did not differ significantly between the sexes (χ^2= 3.73, *DF* = 2, *P* = 0.155).
310 There was no significant difference in tick load between *TLR2* clusters (χ^2 =
311 2.43, *DF* = 2, *P* = 0.296). Including tick load as a covariate in the model of
312 *Borrelia* prevalence (see above) did not change the significant association
313 between *TLR2* clusters and *Borrelia* infection (χ^2 = 7.44, *DF* = 2, *P* = 0.024).
314 A model selection procedure based on AIC_c values largely confirmed these
315 results (except for the sex differences in infection) and revealed that a model
316 containing *TLR2* genotype, age class and the interaction between *TLR2*
317 genotype and age class best explained variation in *Borrelia* infection status of

318 voles. All other candidate models had a $\Delta AIC_c > 2$. Model averaging revealed
319 a significant effect of *TLR2* genotype, with individuals with a $TLR2_{c1}$ / $TLR2_{c2}$
320 (95% CI: -3.57 – -0.33) and a $TLR2_{c2}$ / $TLR2_{c2}$ (95% CI: -6.46 – -0.29)
321 genotype being significantly less likely to be *Borrelia* infected. Furthermore,
322 juveniles (95% CI: -6.67 – -1.22) and subadults (95% CI: -4.44 – -1.82) were
323 significantly less likely to be *Borrelia* infected.
324
325
326 Discussion
327 Using a candidate gene approach we have shown that polymorphisms
328 at the innate immune receptor TLR2 are associated with *B. afzelii* infection
329 status in a natural rodent population. Genetic diversity at *TLR2* was high in
330 the studied bank voles, and a reconstructed haplotype network revealed that
331 *TLR2* variants grouped in well-defined clusters. The same clusters were also
332 found in other bank vole populations in Southern Sweden (0.3-342 km apart)
333 [33], indicating that this unusual haplotype network is not specific to our study
334 population.
335 Homology modelling based on the human TLR2-TLR1 complex
336 indicated that the polymorphism at position 276 (Thr276Ala), which was one
337 of the six linked nonsynonymous mutations that defined the two major
338 haplotype clusters, may have pronounced functional consequences for ligand
339 binding of the TLR2-TLR1 complex [35]. Consistent with the hypothesis that
340 Thr276Ala, or linked amino acid mutations, affect the function of the TLR2-
341 TLR1 complex, we observed marked differences in *Borrelia* prevalence in
342 adult voles with different *TLR2* genotypes. Animals with two haplotypes

343 belonging to cluster $TLR2_{c2}$ were almost three times less likely to be *Borrelia*
344 infected compared to animals with two haplotypes belonging to cluster
345 $TLR2_{c1}$. Whereas there were clear differences in *Borrelia* infection status, we
346 found no difference in tick load between *TLR2* haplotype clusters, suggesting
347 that *TLR2* genotype influences the hosts' resistance to *B. afzelii* rather than
348 their rate of exposure (e.g., through indirect effects of *TLR2* genotype on host
349 behaviour). Adult males were more heavily infected with *B. afzelii* than adult
350 females. This might be due to a higher moving activity of males, which
351 influences *Borrelia* encounter [51], and / or higher testosterone levels, which
352 negatively influences parasite control [52]. Despite these behavioural and
353 physiological differences, the relationship between *TLR2* genotype and *B.*
354 *afzelii* infection was similar in the two sexes. Furthermore, the association
355 between *TLR2* clusters and *B. afzelii* infection was pronounced in adults after
356 dispersal, but absent in juveniles before dispersal [53], indicating that the
357 observed pattern is unlikely due to non-genetic factors shared by family
358 members (i.e. *B. afzelii* abundance in a territory). The conclusion that there is
359 a causal relationship between *TLR2* polymorphisms and the voles' resistance
360 to *B. afzelii* is in line with the results of knock-out studies in laboratory mice,
361 which have identified *TLR2* as a candidate gene for *Borrelia* resistance [23-
362 27]. Nevertheless, given the correlative nature of our study, we cannot
363 exclude the possibility that a linked locus, rather than *TLR2* itself, is driving
364 the observed relationship.
365 What immunological mechanisms could mediate an association
366 between *TLR2* polymorphisms and *B. afzelii* infection status? In principle,
367 improved resistance could be a result of enhanced innate or acquired immune

368 responses. Infections with *B. afzelii* and other Lyme Borreliosis spirochaetes
369 typically result in chronic infections in their natural hosts, with low rates of
370 clearance once the infection has established and disseminated [47]. This
371 suggests that the improved resistance conferred by variants of cluster $TLR2_{c2}$
372 acts via mechanisms expressed early during infection, that is, effectors
373 belonging to the innate immune system. Recent studies have shown that
374 TLRs can activate the complement system [54], a component of innate
375 immunity known to be important for resistance against *Borrelia* [55]. Thus, one
376 possibility is that a higher affinity of cluster $TLR2_{c2}$ haplotypes to *B. afzelii*
377 ligands results in a stronger complement response.
378 Haplotype 6 was the most common *TLR2* haplotype in the host
379 population, but unlike the common haplotype of cluster $TLR2_{c1}$ (haplotype 1),
380 which occurred in six variants, this haplotype has not yet accumulated any
381 synonymous nucleotide substitutions. Synonymous nucleotide substitutions
382 are considered to be selectively neutral (but see [56, 57]), and to accumulate
383 over time at a gene specific point-mutation rate [58]. The complete lack of
384 synonymous nucleotide substitutions in haplotype 6, despite its high
385 frequency, is consistent with recent positive selection that has favoured this
386 haplotype. This is also reflected by significant *Fay and Wu's H* and *Li's MFDM*
387 tests, which both indicate that positive selection has acted on *TLR2* [40, 59].
388 Although the latter test is comparably robust against demographic processes
389 [40], it is important to acknowledge that it is difficult to fully disentangle
390 selection and demography with the currently available data. Nevertheless, the
391 structure of the haplotype network and the results of the neutrality tests in
392 combination with the finding that animals carrying haplotype 6 had lower *B.*

393 *afzelii* prevalence suggests that this *TLR2* variant may have increased in
394 frequency as a result of parasite-mediated selection (by *Borrelia* or other
395 pathogens).
396 While our results are in line with the hypothesis that parasite-mediated
397 selection has shaped *TLR2* evolution in bank voles, it is as yet difficult to
398 assess the role of *B. afzelii* as selective agent because very little is known
399 about the fitness-consequences of *Borrelia* infection in natural populations. In
400 white-footed mice (*Peromyscus leucopus*), experimental infection with *B.*
401 *burgdorferi* sensu stricto in the laboratory led to carditis and multifocal arthritis
402 [60], which likely affects host survival and / or reproduction in the wild.
403 However, the two (correlative) studies performed in wild-living hosts to date,
404 one in white-footed mice [61] and one in black-legged kittiwakes (*Rissa*
405 *tridactyla*) [62], did not find indication for survival costs of *Borrelia* infection.
406 Yet, the strength of selection required to drive a selective sweep is relatively
407 low. For example, Obbard et al (2011) [63] estimated that the selective
408 advantage driving the evolution of *Ago2*, one of the fastest evolving immune
409 genes in *Drosophila,* was a mere 0.5-1%. Clearly, it would be very difficult to
410 detect such low levels of selection in a field study.
411
412 In conclusion, our study shows that polymorphism at *TLR2* is associated with
413 *Borrelia* infection in wild bank voles, one of the main reservoir hosts of *B.*
414 *afzelii* in Europe. This is the first demonstration of an association between
415 *TLR* polymorphism and parasitism in a natural, non-human population.
416 Together with our previous finding that patterns of *TLR2* diversity and
417 population differentiation in bank voles are consistent with local adaptation

418 processes [33], our results highlight the important, but often neglected [64],
419 role of the innate branch of the vertebrate immune system in mediating
420 resistance to pathogens in wildlife. The recent characterization of TLRs in a
421 range of non-model organisms [19, 65] makes these genes suitable
422 candidates for future research on the molecular ecology of resistance to
423 parasites.
424
425
426
427 Acknowledgements
428 We thank Staffan Bensch, Jenny Morger and Erik Postma for discussion and
429 two anonymous reviewers for comments on the manuscript. This work was
430 funded by the Swedish Research Council (grants 621-2206-2876 and 621-
431 2006-4551 to LR and HW), the Crafoord Foundation (grant 20060662 to LR),
432 and the Swiss National Science Foundation (grants PA0033_121466 and
433 PP00P3_128386 to BT).

434 References

435

436 1. Lazzaro, B. P. & Little T. J. 2009 Immunity in a variable world. *Phil.*
437 *Trans. R. Soc. B* **364**(1513), 15-26. (doi:10.1098/rstb.2008.0141).
438 2. Siva-Jothy, M. T. & Skarstein F. 1998 Towards a functional
439 understanding of "good genes". *Ecol. Lett.* **1**(3), 178-185.
440 3. Woolhouse, M. E. J., Webster J. P., Domingo E., Charlesworth B. &
441 Levin B. R. 2002 Biological and biomedical implications of the co-evolution of
442 pathogens and their hosts. *Nat. Genet.* **32**(4), 569-577.
443 4. Allen, D. E. & Little T. J. 2009 Exploring the molecular landscape of
444 host–parasite coevolution. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 169-176.
445 (doi: 10.1101/sqb.2009.74.022).
446 5. Kaslow, R. A., McNicholl J. & Hill A. V. S. 2008 *Genetic Susceptibility*
447 *to Infectious Diseases.* New York: Oxford University Press.
448 6. Turner, A. K., Begon M., Jackson J. A., Bradley J. E. & Paterson S.
449 2011 Genetic diversity in cytokines associated with immune variation and
450 resistance to multiple pathogens in a natural rodent population. *PLoS Genet.*
451 **7**(10), e1002343. doi:10.1371/journal.pgen.1002343.
452 7. Murphy, K., Travers P. & Walport M. 2008 *Janeway's immunobiology*.
453 7th ed. ed. New York: Garland Science.
454 8. Takeda, K., Kaisho T. & Akira S. 2003 Toll-like receptors. *Annu. Rev.*
455 *Immunol.* **21**, 335-376. (doi:10.1146/annurev.immunol.21.120601.141126).
456 9. Akira, S., Uematsu S. & Takeuchi O. 2006 Pathogen recognition and
457 innate immunity. *Cell* **124**(4), 783-801. (doi:10.1016/j.cell.2006.02.015).

458 10. Akira, S. & Takeda K. 2004 Toll-like receptor signalling. *Nat. Rev.*
459 *Immunol.* **4**(7), 499-511. (doi:10.1038/nri1391).
460 11. Medzhitov, R., Preston-Hurlburt P. & Janeway C. A. 1997 A human
461 homologue of the Drosophila Toll protein signals activation of adaptive
462 immunity. *Nature* **388**(6640), 394-397.
463 12. Roach, J. C., Glusman G., Rowen L., Kaur A., Purcell M. K., Smith K.
464 D., Hood L. E. & Aderem A. 2005 The evolution of vertebrate Toll-like
465 receptors. *Proc. Natl. Acad. Sci. USA* **102**(27), 9577-9582.
466 (doi:10.1073/pnas.0502272102).
467 13. Medzhitov, R. 2001 Toll-like receptors and innate immunity. *Nat. Rev.*
468 *Immunol.* **1**(2), 135-145.
469 14. Netea, M. G., Wijmenga C. & O'Neill L. A. J. 2012 Genetic variation in
470 Toll-like receptors and disease susceptibility. *Nat. Immunol.* **13**(6), 535-542.
471 (doi:10.1038/ni.2284).
472 15. Miller, S. I., Ernst R. K. & Bader M. W. 2005 LPS, TLR4 and infectious
473 disease diversity. *Nat. Rev. Microbiol*. **3**(1), 36-46. (doi:10.1038/nrmicro1068).
474 16. Schröder, N. W. & Schumann R. R. 2005 Single nucleotide
475 polymorphisms of Toll-like receptors and susceptibility to infectious disease.
476 *Lancet Infect. Dis.* **5**(3), 156-164. (doi:10.1016/s1473-3099(05)70023-2).
477 17. Piertney, S. B. & Oliver M. K. 2006 The evolutionary ecology of the
478 major histocompatibility complex. *Heredity* **96**(1), 7-21.
479 (doi:10.1038/sj.hdy.6800724).
480 18. Jackson, J. A., Friberg I. M., Bolch L., Lowe A., Ralli C., Harris P. D.,
481 Behnke J. M. & Bradley J. E. 2009 Immunomodulatory parasites and toll-like

receptor-mediated tumour necrosis factor alpha responsiveness in wild mammals. *BMC Biology* **7**. (doi:16 10.1186/1741-7007-7-16).

19. Tschirren, B., Råberg L. & Westerdahl H. 2011 Signatures of selection acting on the innate immunity gene *Toll-like receptor 2* (*TLR2*) during the evolutionary history of rodents. *J. Evol. Biol.* **24**(6), 1232-1240.

20. Kurtenbach, K., Hanincova K., Tsao J. I., Margos G., Fish D. & Ogden N. H. 2006 Fundamental processes in the evolutionary ecology of Lyme borreliosis. *Nat. Rev. Microbiol.* **4**(9), 660-669.

21. Dennis, D. T. & Hayes E. B. 2002 Epidemiology of Lyme Borreliosis. In *Lyme borreliosis: biology, epidemiology and control* (eds. Gray, J., Kahl O., Lane R. S. & Stanek G.), pp. 251-280. Oxford: CABI Publications.

22. Ostfeld, R. S. 2011 *Lyme Disease:The Ecology of a Complex System*. New York: Oxford University Press.

23. Hirschfeld, M., Kirschning C. J., Schwander R., Wesche H., Weis J. H., Wooten R. M. & Weis J. J. 1999 Cutting edge: Inflammatory signalling by *Borrelia burgdorferi* lipoproteins is mediated by Toll-like receptor 2. *J. Immunol.* **163**, 2382-2386.

24. Wooten, R. M., Ma Y., Yoder R. A., Brown J. P., Weis J. H., Zachary J. F., Kirschning C. J. & Weis J. J. 2002 Toll-like receptor 2 is required for innate, but not acquired, host defense to *Borrelia burgdorferi*. *J. Immunol.* **168**(1), 348-355.

25. Alexopoulou, L., Thomas V., Schnare M., Lobet Y., Anguita J., Schoen R. T., Medzhitov R., Fikrig E. & Flavell R. A. 2002 Hyporesponsiveness to vaccination with *Borrelia burgdorferi* OspA in humans and in TLR1- and TLR2-deficient mice. *Nat. Med.* **8**(8), 878-884.

507 26. Roper, R. J., Weis J. J., McCracken B. A., Green C. B., Ma Y., Weber
508 K. S., Fairbairn D., Butterfield R. J., Potter M. R., Zachary J. F., et al. 2001
509 Genetic control of susceptibility to experimental Lyme arthritis is polygenic
510 and exhibits consistent linkage to multiple loci on chromosome 5 in four
511 independent mouse crosses. *Genes Immun.* **2**(7), 388-397.
512 27. Dennis, V. A., Dixit S., O'Brien S. M., Alvarez X., Pahar B. & Philipp M.
513 T. 2009 Live *Borrelia burgdorferi* spirochetes elicit inflammatory mediators
514 from human monocytes via the Toll-like receptor signaling pathway. *Infect.*
515 *Immun.* **77**(3), 1238-1245. (doi:10.1128/iai.01078-08).
516 28. Schröder, N. W. J., Diterich I., Zinke A., Eckert J., Draing C., von Baehr
517 V., Hassler D., Priem S., Hahn K., Michelsen K. S., et al. 2005 Heterozygous
518 Arg753Gln polymorphism of human TLR2 impairs immune activation by
519 *Borrelia burgdorferi* and protects from late stage Lyme disease. *J. Immunol.*
520 **175**(4), 2534-2540.
521 29. Kurtenbach, K., Schäfer S. M., De Michelis S., Etti S. & Sewell H. S.
522 2002 Borrelia burgdorferi sensu lato in the vertebrate host. In *Lyme*
523 *Borreliosis: Biology, Epidemiology and Control* (eds. Gray, J., Kahl O., Lane
524 R. S. & Stanek G.), pp. 117-148, Oxford: CABI Publications.
525 30. Kurtenbach, K., Kampen H., Dizij A., Arndt S., Seitz H. M., Schaible U.
526 E. & Simon M. M. 1995 Infestation of rodents with larval *Ixodes ricinus* (Acari,
527 Ixodidae) is an important factor in the transmission cycle of *Borrelia*
528 *burgdorferi* s.l. in German woodlands. *J. Med. Entomol.* **32**(6), 807-817.
529 31. Wandeler, P., Ravaioli S. R. & Bucher T. B. 2008 Microsatellite DNA
530 markers for the snow vole (*Chionomys nivalis*). *Mol. Ecol. Res.* **8**(3), 637-639.
531 (doi:10.1111/j.1471-8286.2007.02028.x).

532 32. Hellgren, O., Andersson M. & Råberg L. 2011 The genetic structure of
533 *Borrelia afzelii* varies with geographic but not ecological sampling scale. *J.*
534 *Evol. Biol.* **24**(1), 159-167.
535 33. Tschirren, B., Andersson M., Scherman K., Westerdahl H. & Råberg L.
536 2012 Contrasting patterns of diversity and population differentiation at the
537 innate immunity gene *Toll-like receptor 2* (*TLR2*) across populations of two
538 sympatric rodent species. *Evolution* **66**(3), 720-731.
539 34. Gautam, J. K., Ashish, Comeau L. D., Krueger J. K. & Smith M. F.
540 2006 Structural and functional evidence for the role of the TLR2 DD loop in
541 TLR1/TLR2 heterodimerization and signaling. *J. Biol. Chem.* **281**(40), 30132-
542 30142. (doi:10.1074/jbc.M602057200).
543 35. Jin, M. S., Kim S. E., Heo J. Y., Lee M. E., Kim H. M., Paik S. G., Lee
544 H. Y. & Lee J. O. 2007 Crystal structure of the TLR1-TLR2 heterodimer
545 induced by binding of a tri-acylated lipopeptide. *Cell* **130**(6), 1071-1082.
546 (doi:10.1016/j.cell.2007.09.008).
547 36. Drummond, A. J., Ashton B., Cheung M., Heled J., Kearse M., Moir R.,
548 Stones-Havas S., Thierer T. & Wilson A. 2009 Geneious v4.6. (Available from
549 http://www.geneious.com).
550 37. Stephens, M. & Scheet P. 2005 Accounting for decay of linkage
551 disequilibrium in haplotype inference and missing-data imputation. *Am. J.*
552 *Hum. Genet.* **76**(3), 449-462.
553 38. Stephens, M., Smith N. J. & Donnelly P. 2001 A new statistical method
554 for haplotype reconstruction from population data. *Am. J. Hum. Genet.* **68**(4),
555 978-989.

556 39. Fay, J. C. & Wu C. I. 2000 Hitchhiking under positive Darwinian
557 selection. *Genetics* **155**(3), 1405-1413.
558 40. Li, H. P. 2011 A new test for detecting recent positive selection that is
559 free from the confounding impacts of demography. *Mol. Biol. Evol.* **28**(1), 365-
560 375. (doi:10.1093/molbev/msq211).
561 41. Librado, P. & Rozas J. 2009 DnaSP v5: A software for comprehensive
562 analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451-1452.
563 42. Jensen, J. D., Kim Y., DuMont V. B., Aquadro C. F. & Bustamante C.
564 D. 2005 Distinguishing between selective sweeps and demography using
565 DNA polymorphism data. *Genetics* **170**(3), 1401-1410.
566 (doi:10.1534/genetics.104.038224).
567 43. Clement, M., Posada D. & Crandall K. 2000 TCS: a computer program
568 to estimate gene genealogies. *Mol. Ecol.* **9**(10), 1657-1660.
569 44. Thompson, J., Higgins D. & Gibson T. 1994 CLUSTAL W: improving
570 the sensitivity of progressive multiple sequence alignment through sequence
571 weighting, position-specific gap penalties and weight matrix choice. *Nucleic*
572 *Acids Res.* **22**, 4673-4680.
573 45. Eswar, N., Eramian D., Webb B., Shen M. & Sali A. 2008 Protein
574 structure modeling with MODELLER. *Methods Mol Bio* **426**, 145-159.
575 46. Råberg, L. 2012 Infection intensity and infectivity of the tick-borne
576 pathogen *Borrelia afzelii*. *J. Evol. Biol.* **25**(7), 1448-1453.
577 47. Gern, L., Siegenthaler M., Hu C. M., Leubagarcia S., Humair P. F. &
578 Moret J. 1994 Borrelia burgdorferi in rodents (*Apodemus flavicollis* and *A.*
579 *sylvaticus*) - duration and enhancement of infectivity for *Ixodes ricinus* ticks.
580 *Eur. J. Epidemiol.* **10**(1), 75-80. (doi:10.1007/bf01717456).

581 48. Gliwicz, J. 1988 Seasonal dispersal in non-cyclic populations of
582 *Clethrionomys glareolus* and *Apodemus flavicollis Acta Theriol.* **33**(12-25),
583 263-272.
584 49. R Development Core Team. 2011 R: A language and environment for
585 statistical computing. http://www.R-project.org.
586 50. Bates, D., Maechler M. & Bolker B. 2011 lme4: Linear mixed-effects
587 models using S4 classes. (R package version 0.999375-42.
588 51. Kozakiewicz, M., Choluj A. & Kozakiewicz A. 2007 Long-distance
589 movements of individuals in a free-living bank vole population: an important
590 element of male breeding strategy. *Acta Theriol.* **52**(4), 339-348.
591 (doi:10.1007/bf03194231).
592 52. Hughes, V. L. & Randolph S. E. 2001 Testosterone depresses innate
593 and acquired resistance to ticks in natural rodent hosts: A force for
594 aggregated distributions of parasites. *J. Parasitol.* **87**(1), 49-54.
595 (doi:10.1645/0022-3395(2001)087[0049:tdiaar]2.0.co;2).
596 53. Viitala, J., Hakkarainen H. & Ylonen H. 1994 Different dispersal in
597 Clethrionomys and Microtus *Ann. Zool. Fenn.* **31**(4), 411-415.
598 54. Raby, A. C., Holst B., Davies J., Colmont C., Laumonnier Y., Coles B.,
599 Shah S., Hall J., Topley N., Kohl J., et al. 2011 TLR activation enhances C5a-
600 induced pro-inflammatory responses by negatively modulating the second
601 C5a receptor, C5L2. *Eur. J. Immunol.* **41**(9), 2741-2752.
602 (doi:10.1002/eji.201041350).
603 55. Kurtenbach, K., De Michelis S., Etti S., Schafer S. M., Sewell H. S.,
604 Brade V. & Kraiczy P. 2002 Host association of *Borrelia burgdorferi* sensu lato
605 - the key role of host complement. *Trends Microbiol.* **10**(2), 74-79.

606 56. Ingvarsson, P. K. 2010 Natural selection on synonymous and
607 nonsynonymous mutations shapes patterns of polymorphism in *Populus*
608 *tremula*. *Mol. Biol. Evol.* **27**(3), 650-660. (doi:10.1093/molbev/msp255).
609 57. Chamary, J. V. & Hurst L. D. 2005 Evidence for selection on
610 synonymous mutations affecting stability of mRNA secondary structure in
611 mammals. *Genome Biol.* **6**(9). (doi:R75
612 10.1186/gb-2005-6-9-r75).
613 58. Kimura, M. 1983 *The Neutral Theory of Molecular Evolution*.
614 Cambridge, UK: Cambridge University Press.
615 59. Biswas, S. & Akey J. M. 2006 Genomic insights into positive selection.
616 *Trends Genet.* **22**(8), 437-446. (doi:10.1016/j.tig.2006.06.005).
617 60. Moody, K. D., Terwilliger G. A., Hansen G. M. & Barthold S. W. 1994
618 Experimental *Borrelia burgdorferi* infection in *Peromyscus leucopus*. *J. Wildl.*
619 *Dis.* **30**(2), 155-161.
620 61. Hofmeister, E. K., Ellis B. A., Glass G. E. & Childs J. E. 1999
621 Longitudinal study of infection with *Borrelia burgdorferi* in a population of
622 *Peromyscus leucopus* at a Lyme disease-enzootic site in Maryland. *Am. J.*
623 *Trop. Med. Hyg.* **60**(4), 598-609.
624 62. Chambert, T., Staszewski V., Lobato E., Choquet R., Carrie C., McCoy
625 K. D., Tveraa T. & Boulinier T. in press Exposure of black-legged kittiwakes to
626 Lyme disease spirochetes: dynamics of the immune status of adult hosts and
627 effects on their survival. *J. Anim. Ecol.* (doi:10.1111/j.1365-
628 2656.2012.01979.x).
629 63. Obbard, D. J., Jiggins F. M., Bradshaw N. J. & Little T. J. 2011 Recent
630 and recurrent selective sweeps of the antiviral RNAi Gene *Argonaute-2* in

631 three species of *Drosophila*. *Mol. Biol. Evol.* **28**(2), 1043-1056.
632 (doi:10.1093/molbev/msq280).
633 64. Acevedo-Whitehouse, K. & Cunningham A. A. 2006 Is MHC enough for
634 understanding wildlife immunogenetics? *Trends Ecol. Evol.* **21**(8), 433-438.
635 (doi:10.1016/j.tree.2006.05.010).
636 65. Alcaide, M. & Edwards S. V. 2011 Molecular evolution of the Toll-like
637 receptor multigene family in birds. *Mol. Biol. Evol.* **28**(5), 1703-1715.
638 (doi:10.1093/molbev/msq351).
639
640

Tables

Table 1. ***TLR2*** **haplotype frequencies**

Bank vole *TLR2* haplotype frequencies, haplotype cluster and GenBank Accession number. *N* = 1452 haplotypes.

TLR2 haplotype	Frequency	Cluster	GenBank Accession
1 (1_a-1_f)	40.2%	$TLR2_{c1}$	JN674535
2	17.1%	$TLR2_{c1}$	JN674536
3	2.9%	$TLR2_{c1}$	JN674537
4	2.8%	$TLR2_{c1}$	JX014454
6	27.9%	$TLR2_{c2}$	JN674538
7	3.5%	$TLR2_{c2}$	JN674539
8	1.9%	$TLR2_{c2}$	JN674540
9	0.1%	$TLR2_{c2}$	JX014455
10	3.3%	$TLR2_{c3}$	JN674541
14	0.3%	$TLR2_{c1}$	JN674545

Fig. 1 ***TLR2* haplotype network**

Reconstructed bank vole *TLR2* haplotype network based on 726 individuals. Circle sizes reflect the number of haplotype copies found in the population. The circle size for 50, 150 and 350 copies is given as a reference. The numbers along the connecting lines indicate the positions of the (synonymous and non-synonymous) substitutions that separate different haplotypes. Different colours (and numbers) indicate that haplotypes differ at the amino acid level. Haplotypes 1, 2, 3, 4, 14 grouped into one major haplotype cluster (*TLR2*$_{c1;}$ green), haplotypes 6, 7, 8, 9 grouped into a second major haplotype cluster (*TLR2*$_{c2;}$ red). Haplotype 10 formed a third group.

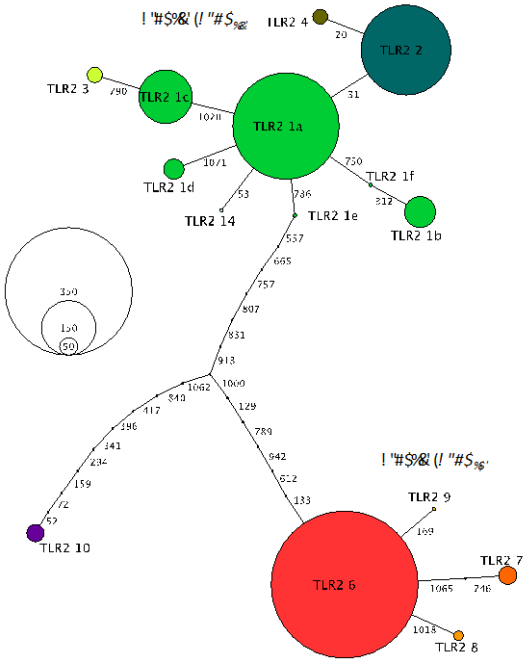

Fig. 2 **Bank vole TLR2-TLR1 complex**

Homology-based structural model of the bank vole TLR2-TLR1 complex showing four amino acid mutations that characterise haplotypes of cluster $TLR2_{c1}$ and $TLR2_{c2}$, respectively. Polymorphic site 276 is located 10.3 Å from the ligand. Amino acid mutations 536 and 565 were outside of the template structure and are not shown. Pink bubbles: polymorphic sites; green: TLR2; gray: TLR1; dusky pink: ligand.

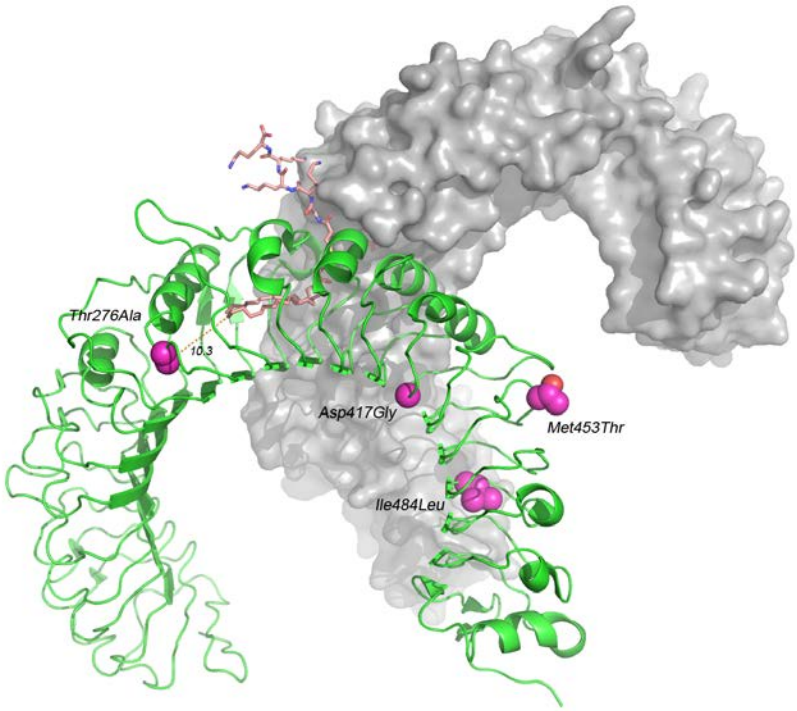

Fig. 3 **Genetic polymorphisms at *TLR2* are associated with *Borrelia* infection status**

Prevalence of *Borrelia* infection in adult bank voles ($N = 234$) with two haplotypes belonging to cluster $TLR2_{c1}$ (c1 / c1), two haplotypes belonging to cluster $TLR2_{c2}$ (c2 / c2), or one haplotype of each cluster (c1 / c2). Mean proportions ± 95% confidence intervals are shown.